ITC 1/55
Information Technology and Control
Vol. 55 / No. 1/ 2026
pp. 222-242
DOI 10.5755/j01.itc.55.1.43005

ID-RTDETR: An Efficient Cross-Scale Fusion Algorithm for Insulator Detection

Received 2025/10/09 | Accepted after revision 2025/12/21

HOW TO CITE: Zhang, X., Chen, H., Zhang, C. (2026). ID-RTDETR: An Efficient Cross-Scale Fusion Algorithm for Insulator Detection. *Information Technology and Control*, 55(1), 222-242. https://doi.org/10.5755/j01.itc.55.1.43005

ID-RTDETR: An Efficient Cross-Scale Fusion Algorithm for Insulator Detection

Xipeng Zhang
School of Mechanical Engineering, Xijing University, Xi'an, 710123, China

Heng Chen*
School of Mechanical Engineering, Xijing University, Xi'an, 710123, China
Engineering Research Center of Hydrogen Energy Equipment & Safety Detection, Xijing University, Xi'an, 710123, China

Chen Zhang
China XD Group XD Baoji Electric Co., Ltd, Baoji, 721306, China

Corresponding author: 18392190260@163.com

Insulator inspection is critical for power system maintenance, but existing methods have challenges in terms of efficiency and accuracy. We propose Insulator Detection Real-Time DEtection TRansformer (ID-RTDETR), an optimized RT-DETR framework with four key innovations for insulator detection. Firstly, Partial Reparameterized Block (PR Block) uses partial convolution with reparameterized in the backbone network, which reduces the computational cost and maintains the feature quality. Secondly, Focusing Diffusion Feature Pyramid Network (FDFPN) realizes the direct interaction of cross-scale features, which can effectively capture different defects. Thirdly, the Attention-based Intra-scale Feature Interaction efficient (AIFI-efficient) reduces the complexity of feature interaction through the additive attention mechanism. Fourth, the Generalized-Inner IoU (G-Inner IoU) loss function is proposed to combine global optimization with local refinement for accurate fault localization. Through the experiment, PR Block reduces parameters by 28.76% and computational cost by 24.78% while improving mAP50:95 by 1.62%. AIFI-Efficient enhances the inference speed of algorithms while maintaining accuracy. FDFPN enhances mAP50:95 by 3.10% through cross-scale feature interaction. G-Inner IoU increased by 2.74% in mAP50:95. The synergistic combination of all four modules achieves optimal performance. After exper-

iments, compared with the baseline RT-DETR, ID-RTDETR improves the mAP50:95 by 1.8 percentage points, reduces the model size by 17.10%, and reaches 71.6 fps. ID-RTDETR provides an efficient and accurate solution for automatic insulator inspection.

KEYWORDS: Insulator Detection, Cross-scale feature fusion, AIFI-Efficient, Additive Attention, Real-time detection.

1. Introduction

Power transmission infrastructure forms the backbone of modern society. Insulators are key components to ensure safe and reliable power transmission. However, insulators exposed for a long time are prone to various defects [38]. Recent studies indicate that 81.3% of power accidents in transmission systems are caused by insulator defects [16]. Failure to detect and address these defects promptly can result in economic losses and safety hazards, making regular inspection essential for grid reliability [1].

The traditional manual detection methods have limitations, including high labor costs, safety risks for personnel working at height, and subjective assessment variability [12]. Recent comprehensive surveys have highlighted both the achievements and limitations of conventional non-destructive testing (NDT) approaches. Traditional radiographic and ultrasonic examination techniques can provide reliable results, but they face challenges in automatic detection and real-time processing [17,27]. Similarly, advanced NDT techniques can now detect small-scale defects with better sensitivity and resolution. However, these techniques still have two main problems: low computational efficiency and the need for expert knowledge to interpret results accurately [25]. These limitations have catalyzed the transition toward machine learning and deep learning approaches. Deep learning can automatically learn defect features from large-scale data and provide fast and consistent detection results [28].

Current deep learning approaches for insulator detection broadly fall into two categories: two-stage detectors that prioritize accuracy and single-stage detectors that emphasize efficiency [8,10,26,39,29]. Two-stage methods, exemplified by the R-CNN family, achieve high detection precision through region proposal networks followed by classification and refinement stages [21]. However, their computational complexity is high. Li et al. [13] indicated that the computational complexity of the two-stage detector stems from its separate region proposal and classification stages. Mittal [20] proposed that two-stage detection poses significant challenges for real-time applications and resource-constrained edge devices. Conversely, single-stage detectors like the YOLO series offer faster inference but often compromise detection accuracy [37]. Xiuling et al. [34] indicated that the single-stage method might compromise the precise localization and classification of challenging targets. Dogra et al. [11] also proposed that the accuracy of single-stage methods limits the development of hybrid architectures.

With the development of automatic insulator detection methods, Liu et al. [18] performed a thorough review of such detection techniques. Their study underscores the shift from conventional image processing methodologies to those grounded in deep learning. The research emphasizes the transformation from traditional image processing techniques to deep learning-based methods. Chen et al. [5] proposed ID-YOLOv7, which addresses the complex backgrounds and subtle defect features in distribution networks. Compared to baseline methods, their approach achieved significant improvements in false detection and omission rates. Similarly, Wang et al. [30] introduced the IF-DETR model, which enhances the detection transformer framework for insulator defect recognition, resulting in an increase in average accuracy by 2.3% over leading existing methods. Xu et al. [35] further advanced the field with MAP-YOLOv8, combining GSConv and SimSPPF modules to reduce computational complexity while improving efficiency. Wang et al. [31] combined the ResNeSt network with a region proposal network, achieving significant improvements in feature representation capabilities and candidate

region generation quality. Zhao et al. [41] optimized the Faster RCNN for improved fault detection and insulator recognition in environments involving transmission lines, concentrating on enhancing the accuracy and robustness of target detection within intricate power settings. Wang et al. [32] proposed a universal plug-in module, MCI-GLA, which significantly improved detection accuracy without significantly increasing computational complexity. Shi et al. [24] based on the latest YOLOv8 architecture, this study conducted an in-depth optimization specifically targeting the small objects.

Carion et al. [4] introduced DEtection Transformer (DETR), marking a paradigm shift in object detection by eliminating the non-maximum suppression (NMS) through end-to-end learning. Despite its innovative design, the computational complexity inherent in the DETR significantly restricts its practical application in real-time scenarios. Recognizing this challenge, Zhao et al. [42] developed the Real-Time DEtection Transformer (RT-DETR), which marks a significant advancement as it is the first object detection system to integrate a real-time, end-to-end methodology grounded in a Transformer architecture. This development opens new avenues for deploying object detection technologies in applications requiring immediate processing and response. While RT-DETR demonstrates improved accuracy-efficiency trade-offs compared to traditional CNN-based detectors, its direct application to insulator detection reveals several limitations. The standard RT-DETR backbone processes all feature channels uniformly, leading to computational redundancy. Additionally, its feature pyramid network struggles with the extreme scale variations present in insulator defects. Although the attention mechanism is powerful, its high computational complexity poses challenges to practical application [14].

Despite extensive research on automated insulator inspection, a critical gap remains in achieving simultaneous improvements in both detection accuracy and computational efficiency. Existing methods typically require choosing between high accuracy with substantial computational costs or real-time performance with compromised detection quality. This trade-off is particularly problematic for edge devices that must maintain real-time performance and high precision [5]. Recent studies have shown that small object detection in high-resolution transmission line images remains a significant challenge [36]. To address these limitations, this study presents Insulator Detection Real-Time DEtection TRansformer (ID-RTDETR). This is an enhanced detection framework specifically designed for efficient and accurate insulator defect identification. Our contributions are as follows:

1. Development of an extensive insulator anomaly detection repository comprising 2,900 meticulously curated images. Data is annotated by professional electrical engineers to ensure high quality. This resource substantially advances computer vision applications within electrical infrastructure inspection domains.
2. This study proposes four innovative modules for insulator detection. Partial Reparameterized Block (PR Block) for efficient feature extraction. Attention-based Intra-scale Feature Interaction Efficient (AIFI) - Efficient facilitating optimized single-scale computation. Feature Direct Fusion Pyramid Network (FDFPN) for cross-scale feature interaction. Generalized-Inner IoU (G-Inner IoU) loss function for precise fault localization combining global optimization with local refinement.
3. This study develops ID-RTDETR, achieving 17.10% model size reduction, 1.80 percentage points mAP50:95 improvement, and 71.6 fps inference speed through comprehensive architectural optimization. ID-RTDETR providing a practical solution for real-world power system inspection.

Through comprehensive experiments on the insulator dataset, we demonstrate that ID-RTDETR achieves advanced performance while significantly reducing computational requirements, providing a practical solution for automated power infrastructure inspection. This article consists of six chapters. Section 1 clearly defines the challenges faced in insulator detection. Section 2 describes repository construction methodology. Section 3 details ID-RTDETR architectural design. Section 4 presents comprehensive experimental validation including ablation studies and comparative analyses. Section 5 examines implications and constraints. Section 6 provides concluding observations.

2. Dataset

This experimental dataset uses a self-created dataset, with images sourced from China XD Group and the internet, as showed in Figure 1. The dataset contains 2,900 original images, including Broken, Flashover, and Insulator. All annotations are carried out by professional electrical engineers. Each image was labeled independently and cross-reviewed by at least two experts. Annotation disagreements were resolved through majority voting. Bounding boxes were drawn tightly around visible defects. This process ensured annotation consistency and minimized subjectivity.

In the subsequent experiments, we divided the dataset into a training set, a validation set and a test set, with a division ratio of 7:2:1, as shown in Table 1.

Table 1
Information about the Insulator detection dataset.

Class	Instances	Broken	Flashover	Insulator
test	291	393	46	254
train	2029	2656	570	1801
val	580	757	166	495
all	2900	3806	782	2550

Specifically, the training set is used to train the model, the validation set is used to adjust the model's hyperparameters during the training process, and the test set is used to evaluate the model's final performance. The division of the dataset ensures the independence between the training data and the test data, prevents overfitting of the model, and also guarantees the reliability of the experimental results. Next, we will introduce the architecture of the RT-DETR algorithm, the process of model training in combination with the dataset, and the specific design.

3. Materials and Methods

3.1. RT-DETR Algorithm

RT-DETR represents a breakthrough in real-time object detection through its innovative three-part framework: a hybrid encoding mechanism, query selection based on IoU awareness, and an adaptable decoding structure.

Central to RT-DETR's design is its hybrid encoder, which diverges from conventional DETR's Transformer-only approach by integrating convolutional and attention-based components. The system leverages backbone networks like ResNet for multi-scale feature extraction, subsequently processing these representations via two specialized modules. The Attention-based Intra-scale Feature Interaction (AIFI) manages feature relationships within individual scales while minimizing computational overhead compared to standard Transformer encoders. Complementing this, the CNN-based Cross-scale Feature Fusion (CCFF) facilitates efficient integration across different feature scales without requiring computationally expensive cross-scale attention operations.

Figure 1
Insulators dataset: (a) Broken; (b) Flashover; (c) Insulator.

(a) (b) (c)

For decoding operations, RT-DETR employs a conventional Transformer decoder architecture that processes encoder outputs alongside IoU-aware queries. Through iterative self-attention and cross-attention operations, the decoder gradually refines object localization and classification predictions.

Notable advantages include reduced dependency on extensive data augmentation and training samples. The framework's flexible decoder layer configuration enables speed optimization without retraining requirements. By eliminating non-maximum suppression while maintaining real-time capabilities, RT-DETR achieves an optimal balance between detection accuracy and processing efficiency. Figure 2 illustrates the complete architectural design of RT-DETR.

Despite its advantages, traditional RT-DETR exhibits several limitations when applied to insulator detection tasks.

1 The standard convolution operations in the backbone network result in high computational redundancy and parameter overhead for processing insulator images.
2 The AIFI module's self-attention mechanism introduces unnecessary computational complexity that hinders real-time performance.
3 The conventional FPN structure lacks direct cross-scale feature interaction, limiting its ability to capture multi-scale insulator defects effectively.
4 The standard IoU loss function struggles with precise localization of insulator faults, lacking the balance between global optimization and local refinement.

Therefore, this paper proposes the ID-RTDETR algorithm, which improves the accuracy and efficiency of insulator detection.

3.2. ID-RTDETR Algorithm

The ID-RTDETR algorithm improves the shortcomings of the traditional RT-DETR algorithm, as shown in Figure 3.

ID-RTDETR first introduces the PRConv (Partial Reparameterized Convolution) module into the backbone network, enhancing feature extraction efficiency through partial convolution and reparameterization techniques. Next, it designs the FDFPN (Focusing Diffusion Feature Pyramid Network), which enables direct interaction between features across layers through the Focus Feature module, breaking the hierarchical limitations of traditional FPN. Next, the AIFI-Efficient module is proposed to make the complexity of mathematical operations decrease of feature

Figure 2
RT-DETR network structure.

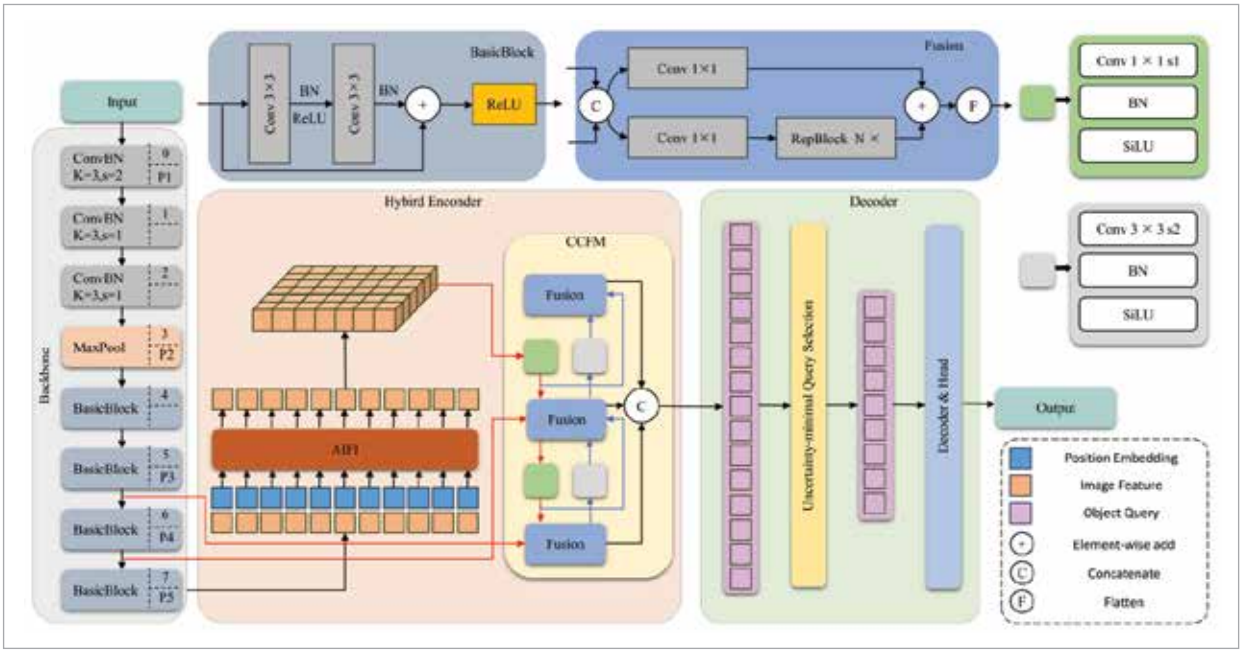
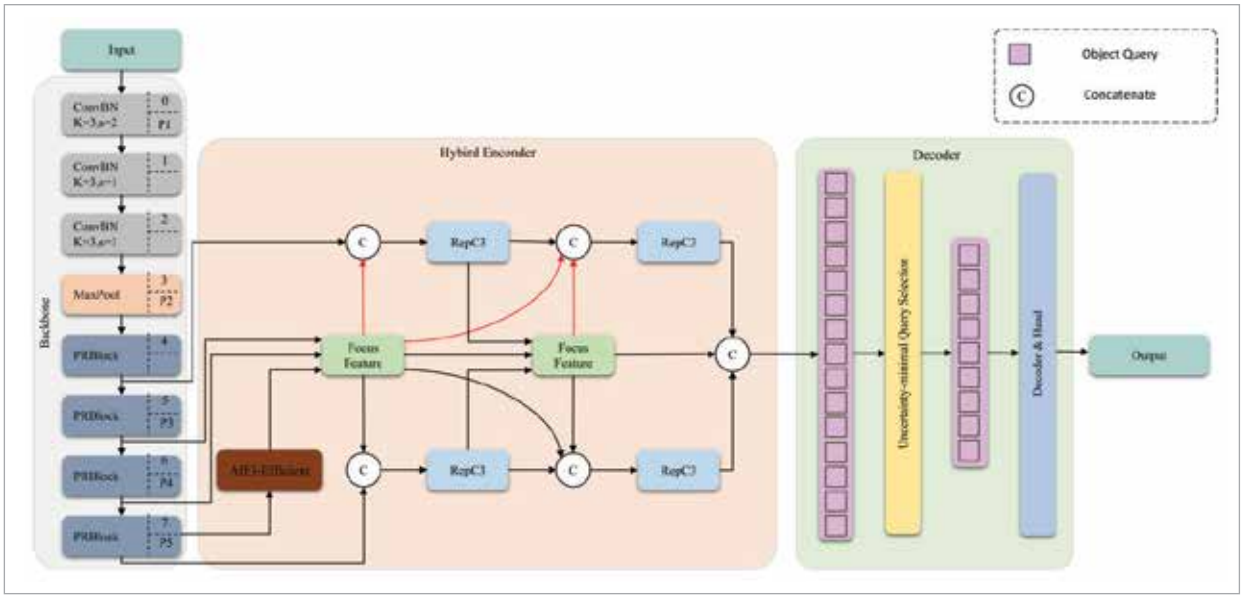


Figure 3
ID-RTDETR network structure.

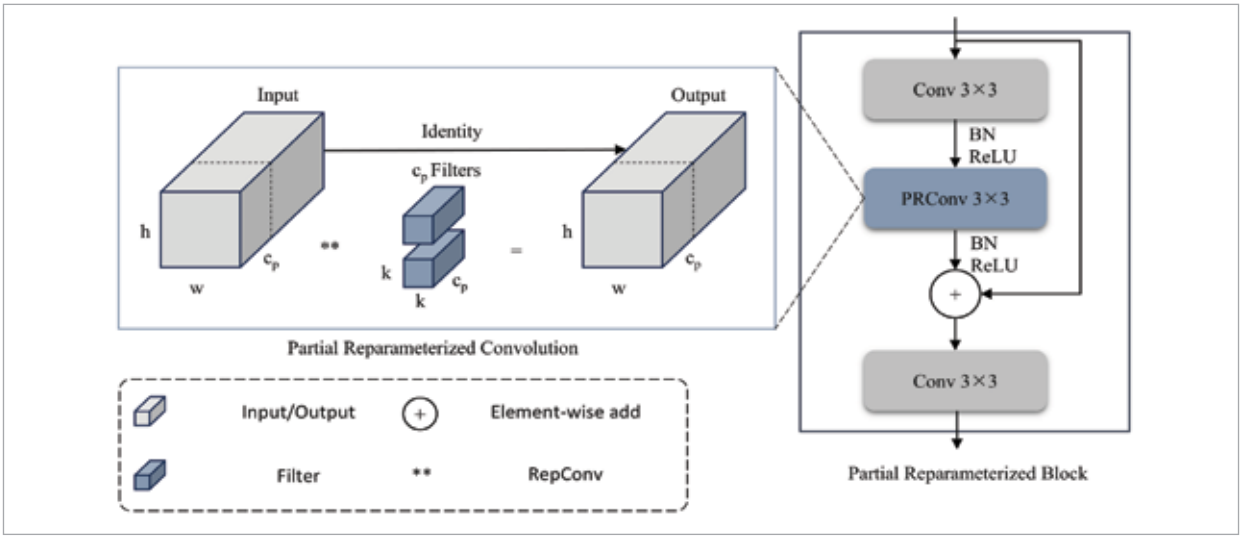


interaction within the same scale. The final stage of the model integrates the Generalized-Inner IoU (G-Inner IoU) to elevate the fidelity of predicted bounding boxes.

3.2.1. Design the PR Block

The traditional BasicBlock structure extracts features through convolution operations on all channels, resulting in computational redundancy and insufficient feature representation capabilities. And traditional convolution processes all channels identically, leading to wasted computational resources and difficulty in effectively capturing key defect features. To address this, we propose an improved structure called PR Block, which combines partial convolution [6] and reparameterization [9] techniques. This module conducts convolution operations exclusively on a selected subset of

Figure 4
Partial Reparameterized Block structure.



the channels from the input feature map, while leaving the other channels intact. This approach effectively strikes a balance between computational efficiency and the preservation of crucial local details within the data. The improved block is shown in Figure 4.

The core workflow of the PRConv module is based on channel separation and selective processing mechanisms. Firstly, the input feature map channels are divided into two parts. The first quarter of the channels undergo convolution operations, while the remaining channels remain unchanged through identity mapping. During the training phase, reparameterized convolution enhances the model's expressive power through a multi-branch structure. Specifically, the input feature map is first divided into multiple subsets, each of which is processed through different convolution operations and optimized through backpropagation. In the reasoning stage, multiple trained branch convolution operations are fused into a 3×3 convolution. This mechanism enables the network to allocate computational resources only to necessary channels, thereby more accurately capturing minor defects and significantly reducing FLOPs.

$$h \times w \times k^2 \times c_p^2 . \quad (1)$$

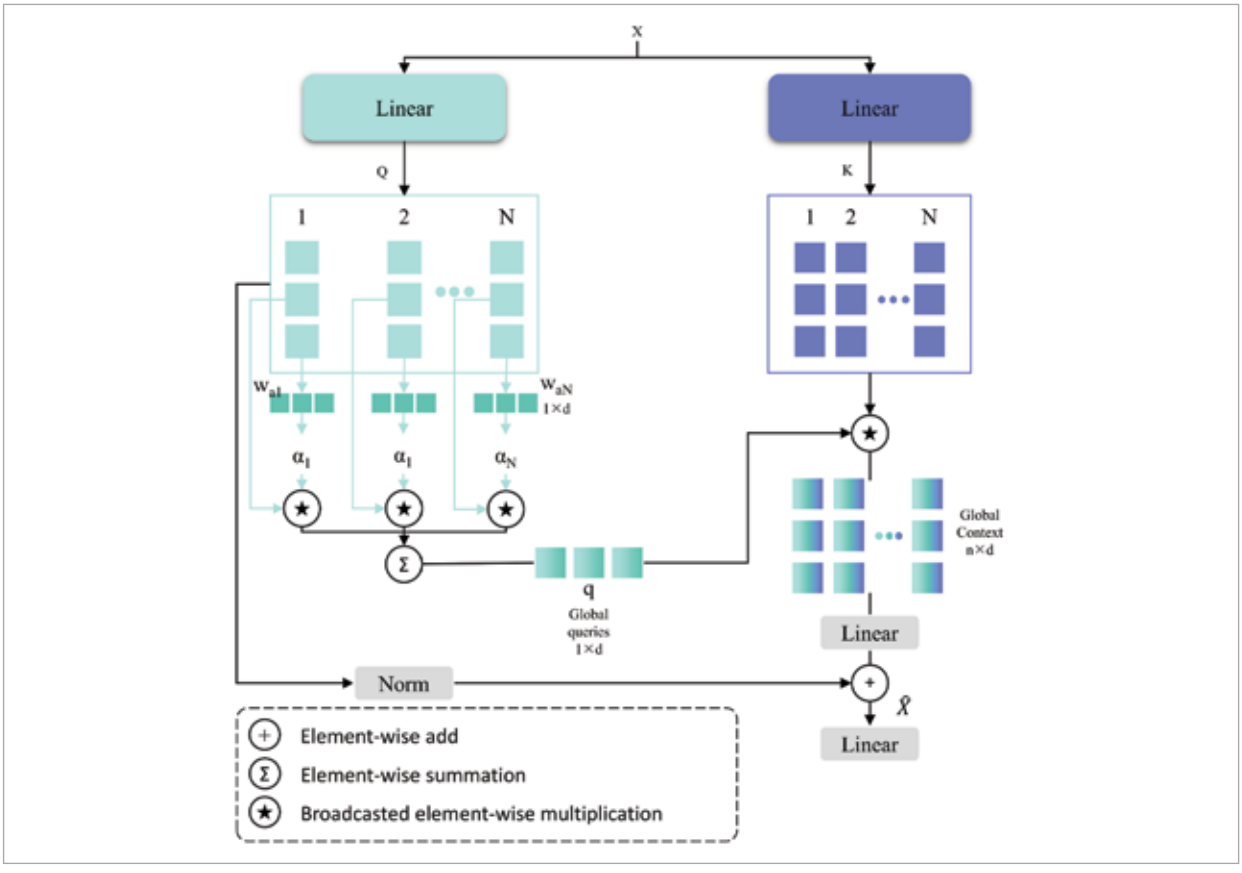
At the same time, partial convolution strategies can effectively reduce the amount of computation. The formula for the amount of computation is:

$$h \times w \times 2c_p + k^2 \times c_p^2 . \quad (2)$$

3.2.2. AIFI-Efficient module

To enhance model capabilities and boost detection accuracy while minimizing missed defects in welding inspections, we developed a component called

Figure 5
Architecture of AIFI-Efficient mechanism.



AIFI-Efficient. Standard self-attention approaches exhibit quadratic computational demands relative to sequence length n, creating substantial processing burdens for high-definition feature representations. Our AIFI module leverages an additive attention strategy that achieves linear complexity, delivering marked improvements in processing speed [23].

The fundamental component within AIFI-Efficient centers on its streamlined Additive Attention architecture, illustrated in Figure 5.

Module input features undergo dual parallel linear projections, yielding query matrix Q and key matrix K, both structured with n×d dimensionality. This dual-pathway architecture preserves query-key interplay foundations.

The system assigns a location-dependent weight vector to each spatial location. These weights are combined with their respective query vectors to produce a location-weighted output:

$$\alpha_i = Q_i \times w_{ai} . \quad (3)$$

An aggregation process then consolidates these position-weighted elements via summation, forming a comprehensive global query representation q that captures sequence-wide query characteristics:

$$q = \sum_{i=1}^{n} \alpha_i . \quad (4)$$

Global contextual computation occurs through element-wise multiplication, distributing aggregated query data across all spatial coordinates to establish comprehensive context encoding. Following this, the contextual representation undergoes linear projection for feature transformation. Concurrently, normalized input features combine with projected context via residual linkage. This intermediate result receives additional linear processing, producing the module's final output representation.

3.2.3. Design the FDFPN

The conventional Feature Pyramid Network employs a simple layer-wise fusion strategy, which struggles to fully leverage the complementary information between features of different scales. Additionally, the unidirectional information flow in FPN limits sufficient interaction between multi-scale features, leading to false positives and false negatives in complex contexts. As shown in Figure 6, FDFPN is proposed, which achieves efficient detection of insulator defects through an innovative feature focusing mechanism and multi-scale diffusion strategy. The structure of FDFPN is shown in Figure 6.

As shown in Figure 7, the core design concept of FDFPN is to break the hierarchical limitations of traditional FPNs and enable direct interaction between features across different scales. The network introduces a Focus Feature module as its basic building block, which can simultaneously process feature maps from three different resolution levels.

Figure 6
Focusing Diffusion Feature Pyramid Network structure.

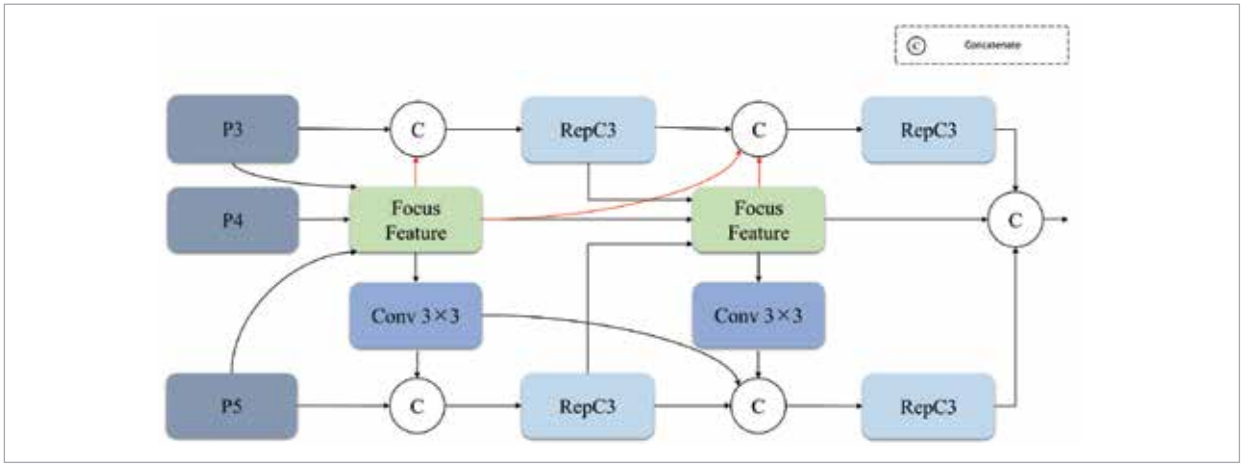
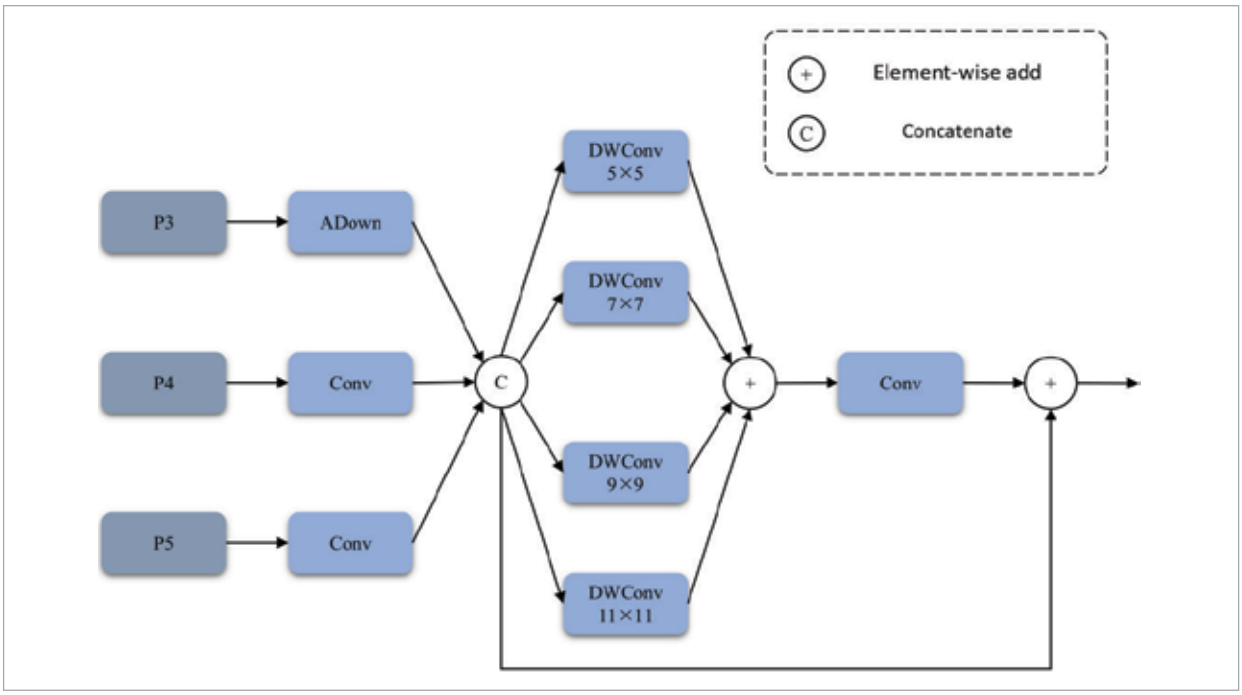


Figure 7
Focus Feature structure.



This module divides the pooled features into two branches: one branch performs downsampling through strided convolution, while the other branch first performs max pooling and then adjusts the number of channels through convolution. This design ensures that useful details and semantic information for defect detection are retained to the greatest extent possible while reducing the resolution.

Upon the conclusion of the scale alignment process, features extracted from three distinct levels are integrated in the channel dimension, resulting in the creation of fused features that encapsulate multi-scale information. This method allows for a richer representation of the data, ensuring that various dimensions of information are considered together. To elevate the network's capacity for detecting defects across a variety of scales, the FDFPN incorporates a multi-scale diffusion mechanism. This approach facilitates the dissemination of information across different scales, further improving the network's sensitivity and accuracy in identifying defects. By leveraging this mechanism, the network can more effectively respond to and analyze defects, regardless of their size or scale. This mechanism uses a set of depth-separable convolutional kernels of different sizes to perform parallel processing on the fused features. Each convolutional kernel is responsible for capturing feature patterns within a specific receptive field range, ensuring that the network can simultaneously respond to both fine cracks and large-area defects.

The overall architecture of FDFPN adopts a two-stage cascaded design, forming a feature pyramid structure with bidirectional information flow. The initial phase of the process begins with the extraction of original multi-scale features from the backbone network. These foundational features are subsequently refined through the first Focus Feature module, which enhances them and generates preliminary enhanced features. Following this enhancement, the features undergo a sequence of upsampling and downsampling operations. This processing allows for the creation of a new, optimized set of multi-scale features that will support further analysis and development in the subsequent stages of the process. The second stage builds upon the first stage by reapplying the Focus Feature module for feature focusing, producing the final detection features.

3.2.4. Loss Function G-Inner IoU

Although the traditional GIoU loss function [40] solves the problem of vanishing gradient of IoU in non-overlapping cases by introducing a minimum bounding rectangle, its ability to optimize the accuracy ofw bounding box regression is still limited. Especially when dealing with subtle defects on the insulator surface, the GIoU loss is difficult to provide an effective gradient signal to further optimize the positioning accuracy in the later training stage when the bounding boxes are already well overlapped. At the same time, the morphological diversity of insulator defects requires the detection algorithm to not only accurately locate the overall area of the defect, but also accurately capture the core part of the defect. To solve this problem, this paper proposes G-Inner IoU loss function, which organically combines the global optimization ability of GIoU and the local refinement mechanism of Inner IoU to achieve more accurate localization of insulator defects.

The operational framework of G-Inner IoU incorporates an embedded auxiliary boundary system designed to strengthen localization precision within target central sectors, whilst maintaining the computational advantages inherent to GIoU. Beginning with the estimated boundary B_p and true boundary B_g, this methodology establishes the fundamental overlap ratio through standard intersection-union computation:

$$\mathrm{IoU}=\frac{\left|B_p\cap B_g\right|}{\left|B_p\cup B_g\right|}. \tag{5}$$

Then we build the inner helper box with a scale of \ alpha, where \alpha is the scaling factor:

$$B_g^{\mathrm{inner}}=\left(x_g,y_g,\alpha\cdot w_g,\alpha\cdot h_g\right) \tag{6}$$

$$B_p^{\mathrm{inner}}=\left(x_p,y_p,\alpha\cdot w_p,\alpha\cdot h_p\right). \tag{7}$$

The intersection and union ratio of the inner helper boxes is calculated as follows:

$$\mathrm{IoU}_{\mathrm{inner}}=\frac{\left|B_p^{\mathrm{inner}}\cap B_g^{\mathrm{inner}}\right|}{\left|B_p^{\mathrm{inner}}\cup B_g^{\mathrm{inner}}\right|}. \tag{8}$$

The core of G-Inner IoU lies in combining the penalty term of GIoU with the internal IoU mechanism. Let C be the smallest enclosing rectangle containing the predicted and true boxes, then G-Inner IoU is defined as follows:

$$\text{G-Inner IoU}=\omega_1\cdot\mathrm{IoU}+\omega_2\cdot\mathrm{IoU}_{\mathrm{inner}}-\frac{\left|C-\left(B_p\cup B_g\right)\right|}{\left|C\right|}, \tag{9}$$

where ω_1 and ω_2 are the weight coefficients. The final G-Inner IoU loss function is expressed as follows:

$$L_{\mathrm{G-Inner\ IoU}}=1-\text{G-Inner IoU}. \tag{10}$$

4. Experiment

4.1. Evaluation Metrics

Insulator detection requires a balance between accuracy and efficiency, while also having resource usage requirements. Consequently, the present investigation employed Precision, mean Average Precision, parameter count, computational complexity in GFLOPs, and frames per second as performance indicators for comprehensive model assessment. The formula is as follows:

$$P=\frac{TP}{TP+FP}100\% \tag{11}$$

$$R=\frac{TP}{TP+FN}100\%. \tag{12}$$

TP is a true example, FP is a false positive example, and FN is a false negative example.

$$\mathrm{AP}=\int_0^1 \mathrm{P}(\mathrm{R})\,\mathrm{dR}, \tag{13}$$

where P(R) denotes the precision as a function of recall R.

In discrete form, the AP can be approximated as the sum over all distinct recall levels:

$$\mathrm{AP}=\sum_{n=1}^{N}\left(R_n-R_{n-1}\right)p_n, \tag{14}$$

where P_n and R_n represent the precision and recall at the n^{th} threshold, respectively, and N is the total number of recall points.

This metric quantitatively measures the detection performance by integrating the trade-off between precision and recall. The mean Average Precision (mAP) is then calculated as the mean of AP across all object categories.

$$\mathrm{mAP}=\frac{1}{\mathrm{N}}\sum_{\mathrm{q}}\mathrm{AP}(\mathrm{q}). \tag{15}$$

4.2. Experimental Environment

Table 2 shows the environmental information of the training process. To ensure the stability and computational efficiency of model training, all input images were uniformly processed using an adaptive scaling strategy to a resolution of 640×640 pixels. The number of epochs is 200.

Table 2
Experimental environment configuration.

Types	Configuration
Ubuntu	20.04
Python	3.8.16
PyTorch	1.13
GPU	RTX 4090
CPU	22 vCPU
CUDA	11.7
batch	16
workers	16
learning rate	1×10^{-4}

4.3. Ablation Experiment

To comprehensively evaluate the contributions of each innovative module in the ID-RTDETR algorithm, we conducted an ablation experiment. The ablation experiment aims to independently evaluate the impact of each module on the overall performance, thereby verifying the role of each module in improving detection accuracy and computational efficiency. Specifically, we selected four key modules for ablation analysis: PR Block, AIFI-Efficient, FDFPN and G-Inner IoU. These modules are respectively responsible for optimizing feature extraction, enhancing computational efficiency, strengthening multi-scale feature fusion, and improving the positioning accuracy of bounding boxes. The following experiments will gradually analyze the function of each module, demonstrate how they work together, and ultimately enhance the performance of the model.

With PR Block architecture deployed, experimental observations revealed that model parameters contracted by 28.76%, processing complexity diminished by 24.78%, while detection precision experienced a positive shift of 2.11%. AIFI-Efficient maintained the parameter quantity while increasing mAP50:95 by 1.22 percentage points. FDFPN achieved the highest individual improvement in mAP50:95, reaching 54.23%. This proves the important role of cross-level feature interaction in improving detection accuracy. The G-Inner IoU loss function increases mAP50:95 from 52.60% to 54.04% while maintaining the original computational complexity.

PR Block and AIFI-Efficient mainly contribute to improving computational efficiency, while FDFPN and G-Inner IoU mainly enhance detection accuracy. The combination of PR Block and AIFI-Efficient achieved a P of 81.12% and a FPS of 75.6, demonstrating outstanding efficiency optimization. The final model combines all four modules, achieving the best balance between efficiency and accuracy. The ablation experiment verified the effectiveness and good synergistic effect of the innovative module, achieving a mAP50:95 of 54.40%, which was 1.80 percentage points higher than the baseline. The algorithm parameters have decreased by 39.24% compared to the baseline.

Table 3
Results of ablation experiments.

Methods	PR Block	AIFI-Efficient	FDFPN	G-Inner IoU	P (%)	mAP50 (%)	mAP50:95 (%)	Params (M)	GFLOPs (G)	FPS (f/s)
base 1					77.09	75.17	52.60	38.6	56.9	71.3
2	√				79.20	74.90	53.45	27.5	42.8	70.0
3		√			77.18	74.93	53.82	38.6	57.2	70.4
4			√		74.60	75.29	54.23	43.1	66.1	67.8
5				√	76.90	74.63	54.04	38.6	56.9	69.3
6	√	√			81.12	73.28	53.20	27.5	43.0	75.6
7	√	√	√		75.26	74.47	54.31	32.0	52.2	71.1
ours 8	√	√	√	√	75.10	74.98	54.40	32.0	52.2	71.6

a The "√" symbol indicates the corresponding module improvement.
b The experiment adopted the same configuration and epochs.

4.4. Verify the Validity by Comparison

4.4.1. Backbone Network Comparison

To verify the performance of backbone improved by PR Block, we use comparative experiments. Table 4 shows the comparison of our backbone with multiple backbones.

Experiments show that our backbone achieves 53.45% on mAP50:95, outperforming other methods. While maintaining high accuracy, it also realizes effective control of the number of parameters and calculation. When benchmarked against RT-DETR-r18, the proposed architecture demonstrates 28.76% fewer parameters and 24.78% lower computational overhead, fulfilling lightweight design requirements. Our backbone achieves an inference speed of 70.0 fps, comparable to RT-DETR-r18's 71.3 fps, and significantly outperforms methods such as EfficientViT (44.7 fps), RepViT (58.0 fps), SwinTransformer (30.4 fps), and ConvNext V2 (58.8 fps) This high-speed inference capability shows that our model is effective.

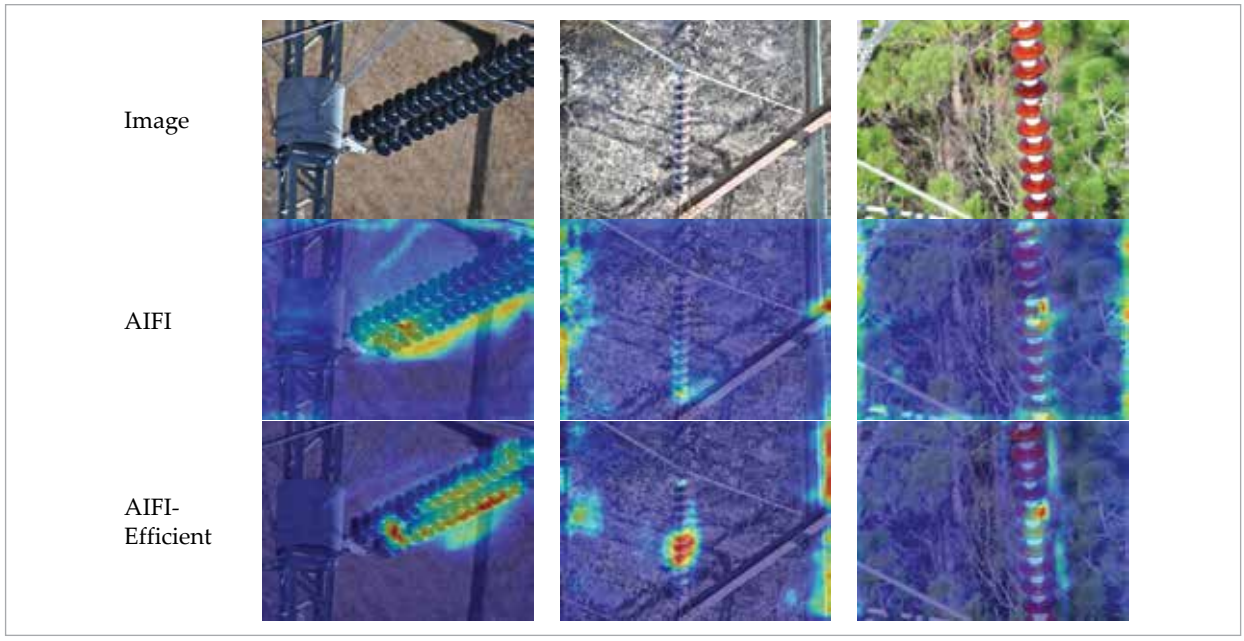
4.4.2. Verify the AIFI-Efficient Module

The Grad-CAM [22] method to visually show the changes in the attention distribution of the model before and after adding the Efficient module. By comparing the detection results of three different sets of scenes, the improvement of model performance by the module is comprehensively evaluated. This is shown in Figure 8.

Table 4
Compare different backbones.

Methods	P (%)	mAP50 (%)	mAP50:95 (%)	Params (M)	GFLOPs (G)	FPS (f/s)
RT-DETR-r18	77.09	75.17	52.60	38.6	56.9	71.3
EfficientViT [19]	75.55	71.49	47.74	21.6	27.2	44.7
RepViT [33]	76.82	70.01	48.69	26.4	36.3	58.0
SwinTransformer [7]	78.40	67.65	47.16	70.2	97.0	30.4
ConvNext V2 [15]	72.98	70.45	48.94	24.1	31.9	58.8
ours	79.20	74.90	53.45	27.5	42.8	70.0

Figure 8
Heat maps generated before and after the addition of AIFI-Efficient.



Examination of gradient-weighted activation mappings illustrates pronounced improvements in spatial attention mechanisms subsequent to Efficient module deployment. The heat map of the original RT-DETR shows that the attention distribution is relatively scattered, while the improved model can locate the core feature area of the target more accurately and reduce the background interference.

The improved model presents more compact and accurate activation regions in the heatmap. Specifically, the compact activation footprint reduces localization ambiguity while strengthening feature-to-object associations. These improvements manifest as decreased false detection rates and elevated recall performance, confirming the Efficient module's effectiveness in optimizing spatial perception capabilities. The architectural enhancement thus establishes more discriminative feature representations, enabling robust object identification across challenging visual conditions.

The quality of heatmap generation is not degraded by the pursuit of computational efficiency. On the contrary, the more focused attention distribution indicates that the module achieves substantial improvement in detection performance while maintaining efficient computation. The heat map comparison experiment fully proves the effectiveness of the AIFI-Efficient module.

4.4.3. Verify the FDFPN Module

In this study, comparative experiments are used to prove the performance of FDFPN, and the experimental results are shown in Table 5.

The experimental results show that FDFPN shows obvious advantages in multiple key indicators. Most notably, on the most stringent evaluation

Table 5
Compare different networks.

Methods	P (%)	mAP50 (%)	mAP50:95 (%)	Params (M)	GFLOPs (G)	FPS (f/s)
RT-DETR-r18	77.09	75.17	52.60	38.6	56.9	71.3
FDFPN	74.60	75.29	54.23	43.1	66.1	67.8

index mAP50:95, FDFPN reaches 54.23%, which is 1.63 percentage points higher than the 52.60% of RT-DETR-r18. This significant improvement fully proves the effectiveness of our proposed multi-scale feature interaction mechanism. On the mAP50 index, FDFPN also achieves 75.29%, which is slightly higher than the 75.17% of the baseline method, indicating that the model maintains a stable performance in routine detection tasks. Although the number of parameters of FDFPN is increased from 38.6M to 43.1M, and the amount of computation is increased from 56.9 GFLOPs to 66.1 GFLOPs, this moderate increase is in return for a significant improvement in detection accuracy. More importantly, FDFPN still maintains the inference speed of 67.8 fps.

4.4.4. Verify the G-Inner IoU

We compare G-Inner IoU with several existing loss functions. The experimental results are shown in Table 6.

Table 6
Compare different loss functions.

Methods	P (%)	mAP50 (%)	mAP50:95 (%)
GIoU	77.09	75.17	52.60
Focaler IoU	74.36	74.90	53.12
G-Inner IoU	76.90	74.63	54.04

G-Inner IoU shows excellent comprehensive performance in the comparison of different IoU loss functions. G-Inner IoU achieves the highest score of 53.12% on the most important mAP50:95 metric. This result indicates that G-Inner IoU is more robust and accurate when dealing with detection tasks with different IoU overlaps.

4.5. Comprehensive Analysis of Model

Table 7 documentation confirms ID-RTDETR's distinguished performance characteristics within insulator detection frameworks, yielding precision rates of 75.10% coupled with recall measurements of 69.78%, whilst average precision at 50% intersection threshold reaches 74.98% and comprehensive average precision across 50-95% thresholds stand at 54.40%. The detection of normal insulators achieved optimal performance, with an accuracy rate as high as 92.08%. This performance is primarily attributed to the FDFPN module effectively capturing the multi-scale structural features of the insulators. The detection performance for damaged defects is relatively lower, with an accuracy rate of 62.58%, and an mAP50 of 65.59%. The primary reason for the higher detection difficulty is damage boundaries are often irregular, complicating precise localization. Flashover defects exhibit unique detection characteristics, with a relatively high accuracy rate of 70.65%, but a recall rate of only 56.02%. This is primarily due to the limited number of flashover samples, which restricts the model's learning capacity. Furthermore, the mAP50:95 for flashover defects is high. Because of the localization accuracy is relatively good.

Meanwhile, to verify that the performance improvement of the proposed algorithm is statistically significant rather than caused by random fluctuations during the training process, a statistical significance test was conducted in this paper. In this paper, the same experimental environment was used for multiple trainings, and the P-value of mAP50:95 was 0.0011, which was less than the significance threshold of 0.05, indicating that the performance difference of the improved model was statistically significant.

Table 7
Results of our model in the dataset.

Class	Instances	P (%)	R (%)	mAP50 (%)	mAP50:95 (%)
all	580	75.10	69.78	74.98	54.40
Broken	579	62.58	57.95	65.59	42.57
Flashover	39	70.65	56.02	64.40	50.58
Insulator	427	92.08	95.35	94.95	70.05

Figure 9 shows several evaluation metrics that reflect the detection performance of the enhanced model. The fluctuation range of each index in the later stage of training is small, which indicates that the model has fully learned the feature representation of insulator defects.

The confusion matrix representation provides critical insights into ID-RTDETR's discriminative capacity for multi-class insulator recognition tasks. Through systematic evaluation of classification outcomes, this analytical framework enables precise quantification of detection accuracy across distinct insulator variants. Figure 10 illustrates the generated confusion matrix, which delineates the algorithm's classification behavior patterns and potential misclassification tendencies.

The algorithm has a high accuracy rate in insulators, damage and flashover. The algorithm performs reli-

Figure 10
The Confusion Matrix of ID-RTDETR.

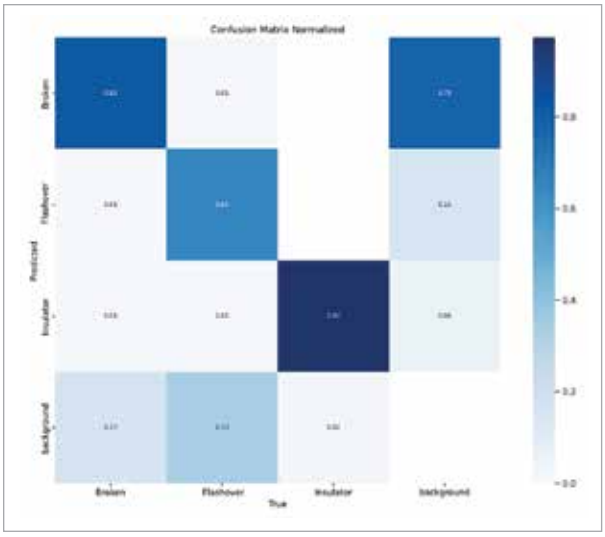
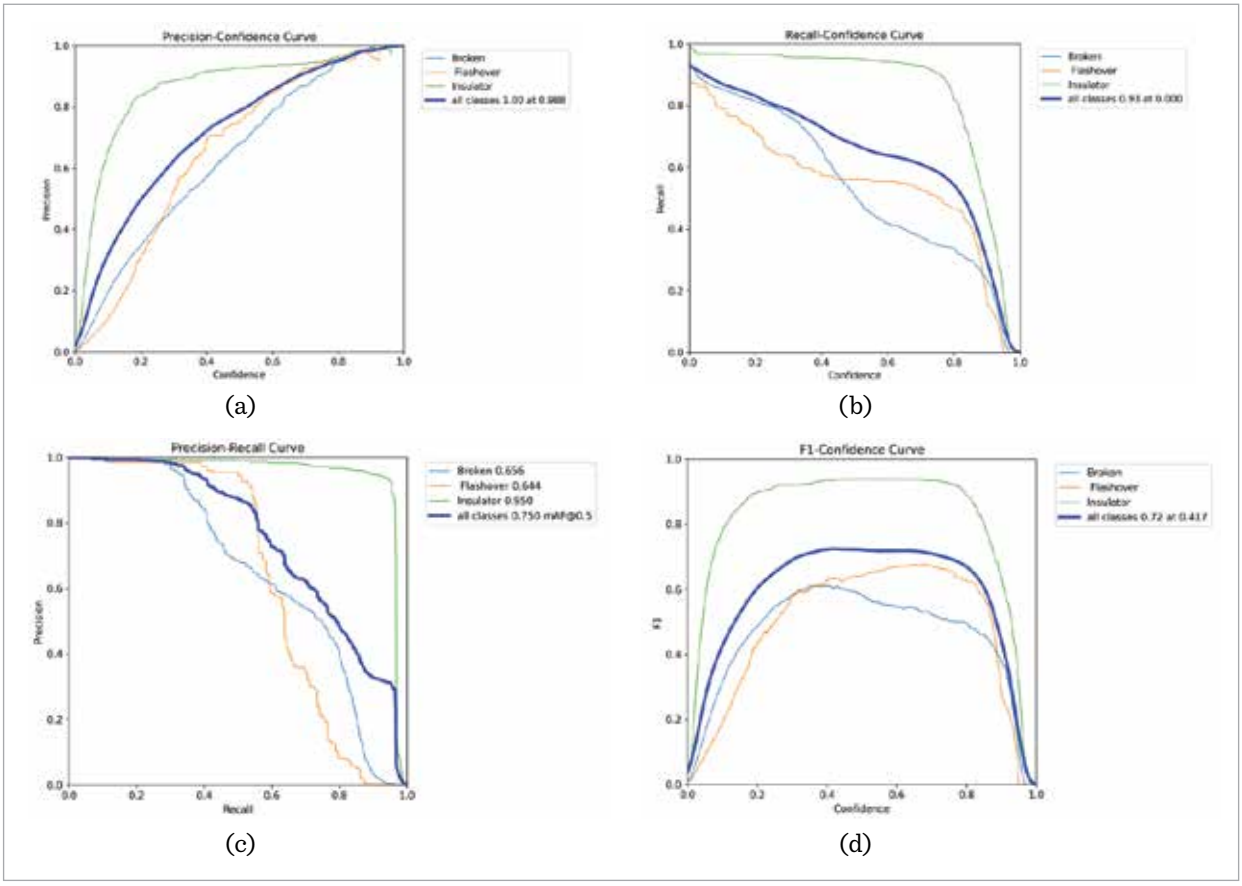


Figure 9
The detection effect of ID-RTDETR: (a) Precision; (b) Recall; (c) mAP50; (d) F1 comparison curve.



ably in key fault categories and can effectively distinguish different types of equipment states.

4.5.1. Comparison of Different Models

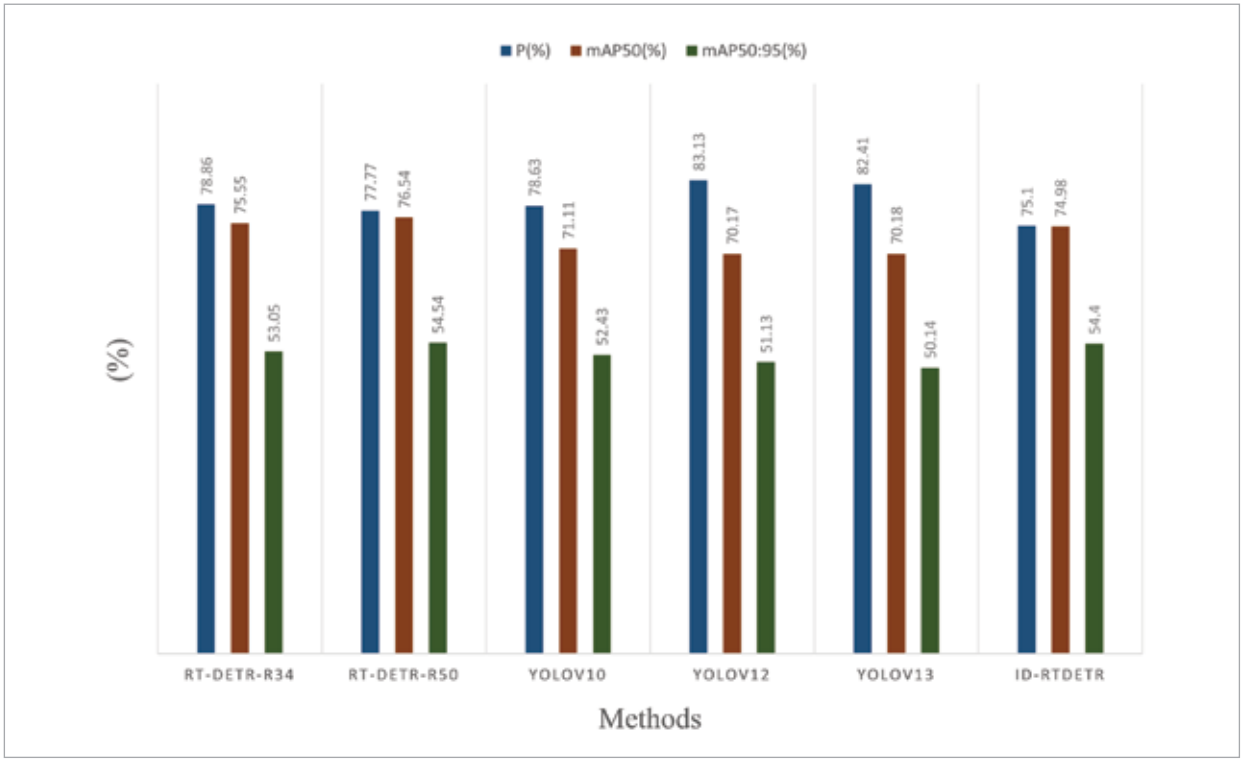
Performance benchmarking of the ID-RTDETR methodology encompassed detailed comparisons with prominent object recognition frameworks currently deployed, featuring RT-DETR implementations (r34 and r50 configurations) together with YOLO advancements (YOLOv10 [2], YOLOv12 [3], and YOLOv13). Corresponding experimental data are visualized in Figure 11 and in Table 8.

ID-RTDETR demonstrates a significant advantage in overall performance. In the most critical mAP50:95 metric, our method achieved 54.40%, ranking first among all comparison algorithms. In terms of model efficiency, ID-RTDETR demonstrates outstanding optimization results. The computational complexity

Table 8
Different algorithm results.

Methods	P (%)	mAP50 (%)	mAP50:95 (%)	Params (M)	GFLOPs (G)	FPS (f/s)
RT-DETR-r34	78.86	75.55	53.05	60.1	88.8	54.4
RT-DETR-r50	77.77	76.54	54.54	82.1	129.6	49.6
YOLOv10	78.63	71.11	52.43	5.8	8.2	123.8
YOLOv12	83.13	70.17	51.13	5.2	5.8	109.6
YOLOv13	82.41	70.18	50.14	5.2	6.2	75.5
ID-RTDETR	75.10	74.98	54.40	32.0	52.2	71.6

Figure 11
Comparing different algorithms.



has been reduced by 41.22% and 59.72% compared to RT-DETR-r34 and r50. Inference speed is a key metric for evaluating the practicality of an algorithm. The ID-RTDETR achieved a real-time processing speed of 71.6 fps, which was significantly outperforms than that of RT-DETR-r34 (54.4 fps) and RT-DETR-r50 (49.6 fps). Although the parameter number and computational complexity of the YOLO series models are relatively small, our method demonstrates a significant advantage in detection accuracy.

It is worth noting that ID-RTDETR successfully breaks the traditional contradiction between accuracy and efficiency. Our algorithm significantly reduces computational costs while improving detection performance. These accurate and rapid features make ID-RTDETR provide a practical solution for insulator detection.

4.5.2. Generalization experiment.

To demonstrate the generalization of the model, this study uses the China Power Line Insulator Dataset (CPLID) public dataset for training. The dataset is shown in Figure 12 and Table 9.

Table 9
Information about the CPLID.

Instances	Insulators	Defect
848	600	248

The training results were compared with those of other models as shown in Table 10.

The detection accuracy and mAP50:95 of ID-RTDETR are superior to several of the most advanced

Table 10
Test results.

Methods	P (%)	mAP50 (%)	mAP50:95 (%)	Params (M)	GFLOPs (G)	FPS (f/s)
RT-DETR	94.55	98.11	78.09	38.6	56.9	72.3
YOLOv12	93.03	96.37	76.52	5.2	5.8	108.2
YOLOv13	94.88	97.25	75.80	17.8	6.2	108.6
ID-RTDETR	96.15	97.90	78.97	32.0	52.2	70.0

Figure 12
Images of the CPLID.

methods at present, reaching 96.15% and 78.97%, respectively. The algorithm is capable of real-time processing and simultaneously maintains a favorable equilibrium between detection accuracy and computational efficiency.

5. Discussion

Our research shows that the insulator detection accuracy of the designed ID-RTDETR algorithm has been improved by 1.69%. The PR Block's partial convolution strategy eliminates redundant computations. AIFI-Efficient reduces complexity from $O(n^2)$ to O(n) via additive attention. FDFPN makes full use of the feature information between different scales through direct cross-scale interaction. G-Inner IoU merges the internal auxiliary box. It effectively handles the irregular defect boundaries in damaged insulators through dual optimization. ID-RTDETR advances automated insulator inspection by achieving real-time, high-accuracy fault detection. Beyond insulator inspection, ID-RTDETR can adapt to the challenges of other scenarios involving small object detection and real-time processing. Because of the efficient feature extraction and multi-scale fusion strategy.

Despite these advances, several limitations warrant discussion. First, extreme weather scenarios could challenge the model's robustness. Secondly, the recall rate of flashover defects is relatively low, which indicates that rare defect types remain challenging. This might be due to insufficient training samples and unique morphological features. Future work should overcome the limitations that have been identified in several ways. Synthetic augmentation methods to augment the dataset could help achieve better performance on uncommon defect types. Researching adaptive attention mechanisms would be able to better detect morphologically varied damages. Further reducing model complexity without compromising accuracy could help enhance algorithmic performance.

6. Conclusions

The ID-RTDETR algorithm represents a step forward in addressing the challenge of balancing computational efficiency and detection accuracy in insulator detection. By integrating architectural innovations such as the PR Block, AIFI-Efficient, FDFPN, and G-Inner IoU, the algorithm achieves high accuracy in detecting insulator defects while maintaining real-time processing speeds. This research is conducive to the development of practical and efficient automatic insulator detection solutions, which will promote real-time detection tasks in various industrial applications.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, author-ship, and/or publication of this article.

Data Sharing Agreement

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

1. Ahmed, F., Mohanta, J. C., Keshari, A. Power Transmission Line Inspections: Methods, Challenges, Current Status and Usage of Unmanned Aerial Systems. Journal of Intelligent & Robotic Systems, 2024, 110(2), 54. https://doi.org/10.1007/s10846-024-02061-y
2. Alkhammash, E. H. A Comparative Analysis of YOLOv9, YOLOv10, YOLOv11 for Smoke and Fire Detection. Fire (2571-6255), 2025, 8(1). https://doi.org/10.3390/fire8010026
3. Buleu, B., Robu, R., Filip, I. A Deep Learning-Based System for Automatic License Plate

Recognition Using YOLOv12 And PaddleOCR. Applied Sciences (2076-3417), 2025, 15(14). https://doi.org/10.3390/app15147833

4. Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A. End-To-End Object Detection with Transformers. European Conference on Computer Vision. Cham: Springer International Publishing, 2020, 213-229. https://doi.org/10.1007/978-3-030-58452-8_13
5. Chen, B., Zhang, W., Wu, W., Li, Y., Chen, Z., Li, C. ID-YOLOv7: An Efficient Method for Insulator Defect Detection in Power Distribution Network. Frontiers In Neurorobotics, 2024, 17, 1331427. https://doi.org/10.3389/fnbot.2023.1331427
6. Chen, J., Kao, S., He, H., Zhuo, W., Wen, S., Lee, C., Chan, S. Run, Don't Walk: Chasing Higher FLOPS For Faster Neural Networks. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023, 12021-12031. https://doi.org/10.1109/CVPR52729.2023.01157
7. Choi, S., Kim, S., Jung, H. Optimized Faster R-CNN With SwinTransformer for Robust Multi-Class Wildfire Detection. Fire, 2025, 8(5), 180. https://doi.org/10.3390/fire8050180
8. Demetriou, D., Mavromatidis, P., Robert, P. M., Papadopoulos, H., Petrou, M. F., Nicolaides, D. Real-Time Construction Demolition Waste Detection Using State-Of-The-Art Deep Learning Methods; Single-Stage Vs Two-Stage Detectors. Waste Management, 2023, 167, 194-203. https://doi.org/10.1016/j.wasman.2023.05.039
9. Ding, X., Zhang, X., Ma, N., Han, J., Ding, G., Sun, J. RepVGG: Making VGG-Style ConvNets Great Again. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, 13733-13742. https://doi.org/10.1109/CVPR46437.2021.01352
10. Diwan, T., Anirudh, G., Tembhurne, J. V. Object Detection Using YOLO: Challenges, Architectural Successors, Datasets and Applications. Multimedia Tools and Applications, 2023, 82(6), 9243-9275. https://doi.org/10.1007/s11042-022-13644-y
11. Dogra, A. K., Sharma, V., Sohal, H. A Survey of Deep Learning Techniques for Detecting and Recognizing Objects in Complex Environments. Computer Science Review, 2024, 54, 100686. https://doi.org/10.1016/j.cosrev.2024.100686
12. Faisal, M. A. A., Mecheter, I., Qiblawey, Y., Fernandez, J., Chowdhury, M., Kiranyaz, S. Deep Learning in Automated Power Line Inspection: A Review. Applied Energy, 2025, 385, 125507. https://doi.org/10.1016/j.apenergy.2025.125507
13. Li, C., Huang, H., Mo, H., Wen, Z., Zhou, S., Zhu, Z. Novel Efficient Steel Surface Defect Detection Model Based on ConvNeXt V2 and Squeeze Aggregated Excitation Attention. IEEE Access, 2025, 13, 107512-107531. https://doi.org/10.1109/ACCESS.2025.3581313
14. Li, Z., Dong, Y., Shen, L., Liu, Y., Pei, Y. Development and Challenges of Object Detection: A Survey. Neurocomputing, 2024, 598, 128102. https://doi.org/10.1016/j.neucom.2024.128102
15. Li, Z., Jiang, C., Li, Z. An Insulator Location and Defect Detection Method Based on Improved YOLOv8. IEEE Access, 2024, 12, 106781-106792. https://doi.org/10.1109/ACCESS.2024.3436919
16. Liu, J., Hu, M. M., Dong, J. Y., Lu, X. Summary of Insulator Defect Detection Based on Deep Learning. Electric Power Systems Research, 2023, 224, 109688. https://doi.org/10.1016/j.epsr.2023.109688
17. Liu, T., Zheng, P., Bao, J., Chen, H. A State-Of-The-Art Survey of Welding Radiographic Image Analysis: Challenges, Technologies and Applications. Measurement, 2023, 214, 112821. https://doi.org/10.1016/j.measurement.2023.112821
18. Liu, X., Peng, H., Zheng, N., Zong, M., Zhu, J. EfficientViT: Memory Efficient Vision Transformer with Cascaded Group Attention. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023, 14420-14430. https://doi.org/10.1109/CVPR52729.2023.01386
19. Liu, Y., Liu, D., Huang, X., Li, C. Insulator Defect Detection with Deep Learning: A Survey. IET Generation, Transmission & Distribution, 2023, 17(16), 3541-3558. https://doi.org/10.1049/gtd2.12916
20. Mittal, P. A Comprehensive Survey of Deep Learning-Based Lightweight Object Detection Models for Edge Devices. Artificial Intel-

ligence Review, 2024, 57(9), 242. https://doi.org/10.1007/s10462-024-10877-1

21. Safonova, A., Ghazaryan, G., Stiller, S., Main-Knorn, M., Nendel, C., Ryo, M. Ten Deep Learning Techniques to Address Small Data Problems with Remote Sensing. International Journal of Applied Earth Observation and Geoinformation, 2023, 125, 103569. https://doi.org/10.1016/j.jag.2023.103569
22. Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., Batra, D. Grad-CAM: Visual Explanations from Deep Networks Via Gradient-Based Localization. Proceedings of the IEEE International Conference on Computer Vision. 2017, 618-626. https://doi.org/10.1109/ICCV.2017.74
23. Shaker, A., Maaz, M., Rasheed, H., Khan, S., Yang, M., Khan, F. SwiftFormer: Efficient Additive Attention for Transformer-Based Real-Time Mobile Vision Applications. Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023, 17425-17436. https://doi.org/10.1109/ICCV51070.2023.01598
24. Shi, W., Lyu, X., Han, L. SONet: A Small Object Detection Network for Power Line Inspection Based on YOLOv8. IEEE Transactions on Power Delivery, 2024, 39(5), 2973-2984. https://doi.org/10.1109/TPWRD.2024.3450185
25. Silva, M. I., Malitckii, E., Santos, T. G., Vilaça, P. Review of Conventional and Advanced Non-Destructive Testing Techniques for Detection and Characterization of Small-Scale Defects. Progress In Materials Science, 2023, 138, 101155. https://doi.org/10.1016/j.pmatsci.2023.101155
26. Sirisha, U., Praveen, S. P., Srinivasu, P. N., Barsocchi, P., Bhoi, A. K. Statistical Analysis of Design Aspects of Various YOLO-Based Deep Learning Models for Object Detection. International Journal of Computational Intelligence Systems, 2023, 16(1), 126. https://doi.org/10.1007/s44196-023-00302-w
27. Sun, H., Ramuhalli, P., Jacob, R. E. Machine Learning for Ultrasonic Nondestructive Examination of Welding Defects: A Systematic Review. Ultrasonics, 2023, 127, 106854. https://doi.org/10.1016/j.ultras.2022.106854
28. Tian, C., Zhang, X., Liang, X., Li, B., Sun, Y., Zhang, S. Knowledge Distillation with Fast CNN For License Plate Detection. IEEE Transactions on Intelligent Vehicles, 2023, 1-7. https://doi.org/10.1109/TIV.2023.3330164
29. Vijayakumar, A., Vairavasundaram, S. YOLO-Based Object Detection Models: A Review and Its Applications. Multimedia Tools and Applications, 2024, 83(35), 83535-83574. https://doi.org/10.1007/s11042-024-18872-y
30. Wang, A., Chen, H., Lin, Z., Han, J., Di, G. RepViT: Revisiting Mobile CNN From ViT Perspective. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024, 15909-15920. https://doi.org/10.1109/CVPR52733.2024.01506
31. Wang, S., Liu, Y., Qing, Y., Wang, C., Lan, T., Yao, R. Detection of Insulator Defects with Improved ResNeSt and Region Proposal Network. IEEE Access, 2020, 8, 184841-184850. https://doi.org/10.1109/ACCESS.2020.3029857
32. Wang, X., Yang, T., Zou, Y. Enhancing Grid Reliability Through Advanced Insulator Defect Identification. PLOS One, 2024, 19(9), e0307684. https://doi.org/10.1371/journal.pone.0307684
33. Wang, Y., Song, X., Feng, L., Zhai, Y., Zhao, Z., Zhang, S., Wang, Q. MCI-GLA Plug-In Suitable for YOLO Series Models for Transmission Line Insulator Defect Detection. IEEE Transactions on Instrumentation and Measurement, 2024, 73, 1-12. https://doi.org/10.1109/TIM.2024.3385817
34. Xiuling, Z., Huijuan, W., Yu, S., Chen, G., Zou, S., Yuan, Q. Starting from the Structure: A Review of Small Object Detection Based on Deep Learning. Image and Vision Computing, 2024, 146, 105054. https://doi.org/10.1016/j.imavis.2024.105054
35. Xu, Z., Tang, X. Transmission Line Insulator Defect Detection Algorithm Based On MAP-YOLOv8. Scientific Reports, 2025, 15(1), 10288. https://doi.org/10.1038/s41598-025-92445-3
36. Yan, X., Wang, W., Lu, F., Fan, H., Wu, B., Yu, J. GFRF R-CNN: Object Detection Algorithm for Transmission Lines. Computers, Materials & Continua, 2025, 82(1). https://doi.org/10.32604/cmc.2024.057797

37. Yuzkat, M., Ilhan, H. O., Aydin, N. Detection of Sperm Cells by Single-Stage and Two-Stage Deep Object Detectors. Biomedical Signal Processing and Control, 2023, 83, 104630.https://doi.org/10.1016/j.bspc.2023.104630

38. Zhang, C., Chang, C., Jamshidi, M. Concrete Bridge Surface Damage Detection Using a Single-Stage Detector. Computer-Aided Civil and Infrastructure Engineering, 2020, 35(4), 389-409. https://doi.org/10.1111/mice.12500

39. Zhang, Q., Liu, Y., Zhang, Y., Zong, M., Zhu, J. Improved YOLOv3 Integrating SENet and Optimized GIoU Loss for Occluded Pedestrian Detection. Sensors, 2023, 23(22), 9089. https://doi.org/10.3390/s23229089

40. Zhang, Q., Zhang, J., Li, Y., Zhu, C., Wang, G. IL-YOLO: An Efficient Detection Algorithm for Insulator Defects in Complex Backgrounds of Transmission Lines. IEEE Access, 2024, 12, 14532-14546.https://doi.org/10.1109/ACCESS.2024.3358205

41. Zhao, W., Xu, M., Cheng, X., Zhao, Z. An Insulator in Transmission Lines Recognition and Fault Detection Model Based on Improved Faster R-CNN. IEEE Transactions on Instrumentation and Measurement, 2021, 70, 1-8. https://doi.org/10.1109/TIM.2021.3112227

42. Zhao, Y., Lv, W., Xu, S., Wei, J., Wang, G. DETRs Beat YOLOs On Real-Time Object Detection. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024, 16965-16974. https://doi.org/10.1109/CVPR52733.2024.01605